ABSTRACT

In accordance with our invention, for two mixture-type probability distribution functions (PDF's), G, H,

$$G(x) = \sum_{i=1}^{N} \mu_i g_i(x), \qquad H(x) = \sum_{k=1}^{K} \gamma_k h_k(x),$$

where G is a mixture of N component PDF's $g_i(x)$, H is a mixture of K component PDF's $h_k(x)$, μ_i and γ_k are corresponding weights that satisfy

$$\sum_{i=1}^{N} \mu_i = 1 \ \ and \ \sum_{k=1}^{K} \gamma_k = 1;$$

we define their distance, $D_M(G, H)$, as

$$D_M(G,H) = \min_{w=[\omega_{ik}]} \sum_{i=1}^{N} \sum_{k=1}^{K} \omega_{ik} d(g_i, h_k)$$

where $d(g_i, h_k)$ is the element distance between component PDF's g_i and h_k and w satisfie

$$\omega_{ik} \geq 0, \ \ 1 \leq i \leq N, 1 \leq k \leq K;$$

and

$$\sum_{k=1}^{K} \omega_{ik} = \mu_i, \ 1 \leq i \leq N, \ \sum_{i=1}^{N} \omega_{ik} = \gamma_k, \ 1 \leq k \leq K.$$

The application of this definition of distance to various sets of real world data is demonstrated.